RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _____10/550, 651_____

Source: _____PCT_____

Date Processed by STIC: _____10/06/2005_____

ENTERED

RAW SEQUENCE LISTING   DATE: 10/06/2005
PATENT APPLICATION: US/10/550,651   TIME: 11:44:04

Input Set : **A:\67675-06 Seq List.txt**
Output Set: **N:\CRF4\10062005\J550651.raw**

```
  3 <110> APPLICANT: THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS
  4       REPRESENTED BY THE SECRETARY OF THE DEPARTMENT OF HEALTH AND
  5       HUMAN SERVICES, CENTERS FOR DISEASE CONTROL AND PREVENTION
  6       Lal, Renu B.
  7       Owen, Sherry M.
  9 <120> TITLE OF INVENTION: IMMUNOGENIC HIV-1 MULTI-CLADE, MULTIVALENT CONSTRUCTS AND
METHODS
 10       OF THEIR USE
 12 <130> FILE REFERENCE: 6395-67675-06
C--> 14 <140> CURRENT APPLICATION NUMBER: US/10/550,651
C--> 14 <141> CURRENT FILING DATE: 2005-09-26
 14 <150> PRIOR APPLICATION NUMBER: PCT/US2004/009767
 15 <151> PRIOR FILING DATE: 2004-03-24
 17 <150> PRIOR APPLICATION NUMBER: US 60/458,880
 18 <151> PRIOR FILING DATE: 2003-03-28
 20 <160> NUMBER OF SEQ ID NOS: 64
 22 <170> SOFTWARE: PatentIn version 3.2
 24 <210> SEQ ID NO: 1
 25 <211> LENGTH: 1557
 26 <212> TYPE: DNA
 27 <213> ORGANISM: Artificial Sequence
 29 <220> FEATURE:
 30 <223> OTHER INFORMATION: Construct encoding polyepitope polypeptide.
 33 <220> FEATURE:
 34 <221> NAME/KEY: CDS
 35 <222> LOCATION: (12)..(1547)
 36 <223> OTHER INFORMATION: Sequence encoding MCMVCTL-ubiquitin polyepitope polypeptide.
 38 <400> SEQUENCE: 1
 39 ctaggctagc t atg cag atc ttc gtg aaa acc ctt acc ggc aag acc atc        50
 40             Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile
 41               1               5                  10
 43 acc ctt gag gtg gag ccc agt gac acc atc gaa aat gtg aag gcc aag        98
 44 Thr Leu Glu Val Glu Pro Ser Asp Thr Ile Glu Asn Val Lys Ala Lys
 45    15               20                  25
 47 atc cag gat aag gaa ggc att ccc ccc gac cag cag agg ctc atc ttt       146
 48 Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe
 49 30               35                  40                  45
 51 gca ggc aag cag ctg gaa gat ggc cgt act ctt tct gac tac aac atc       194
 52 Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile
 53               50                  55                  60
 55 cag aag gag tcg acc ctg cac ctg gtc ctg cgt ctg aga ggt gct gag       242
 56 Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Ala Glu
 57               65                  70                  75
 59 ctc cgc tcc ctc tac aac acc gtg gcc acc ctc tac tgc gtg cac cag       290
```

```
60 Leu Arg Ser Leu Tyr Asn Thr Val Ala Thr Leu Tyr Cys Val His Gln
61          80                  85                  90
63 cgc atc aag atc cgc ctg cgc ccc ggc ggc aag aag aag tac tgg gcc      338
64 Arg Ile Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Trp Ala
65      95                 100                 105
67 agc cgc gag ctg gag cgc ttc aag gcc gcc atc agc ccc cgc acc ctg      386
68 Ser Arg Glu Leu Glu Arg Phe Lys Ala Ala Ile Ser Pro Arg Thr Leu
69 110                 115                 120                 125
71 aac gcc tgg gtg aag gtg gtg aag gcc ttc agc ccc gag gtg atc ccc      434
72 Asn Ala Trp Val Lys Val Val Lys Ala Phe Ser Pro Glu Val Ile Pro
73              130                 135                 140
75 atg ttc agc gcc ctg agc gag ggc gcc acc ccc cag gac ctg aac acc      482
76 Met Phe Ser Ala Leu Ser Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr
77              145                 150                 155
79 atg acc agc acc ctg cag gag cag atc ggc tgg aag gcc gcc aac ccc      530
80 Met Thr Ser Thr Leu Gln Glu Gln Ile Gly Trp Lys Ala Ala Asn Pro
81              160                 165                 170
83 ccc atc ccc gtg ggc gac atc tac aag cgc tgg atc atc ctg ggc ctg      578
84 Pro Ile Pro Val Gly Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu
85      175                 180                 185
87 aac aag atc gtg cgc atg tac agc ccc acc agc atc ttc cgc gac tac      626
88 Asn Lys Ile Val Arg Met Tyr Ser Pro Thr Ser Ile Phe Arg Asp Tyr
89 190                 195                 200                 205
91 gtg gac cgc ttc tac aag acc ctg cgc gcc gtg cag aac gcc aac ccc      674
92 Val Asp Arg Phe Tyr Lys Thr Leu Arg Ala Val Gln Asn Ala Asn Pro
93              210                 215                 220
95 gac tgc aag acc atc ctg aag gcc ctg gcc tgc cag ggc gtg ggc ggc      722
96 Asp Cys Lys Thr Ile Leu Lys Ala Leu Ala Cys Gln Gly Val Gly Gly
97              225                 230                 235
99 ccc ggc cac aag aag gcc gcc atc acc ctg tgg cag cgc ccc ctg gtg      770
100 Pro Gly His Lys Lys Ala Ala Ile Thr Leu Trp Gln Arg Pro Leu Val
101             240                 245                 250
103 acc gtg ctg gac gtg ggc gac gcc tac ttc agc gtg tgg aag ggc agc      818
104 Thr Val Leu Asp Val Gly Asp Ala Tyr Phe Ser Val Trp Lys Gly Ser
105     255                 260                 265
107 ccc gcc atc ttc cag agc aag ctt cgc ggc ccc ggc cgc gcc ttc gtg      866
108 Pro Ala Ile Phe Gln Ser Lys Leu Arg Gly Pro Gly Arg Ala Phe Val
109 270                 275                 280                 285
111 acc atc aag gcc gcc gcc tgc acc ccc tac gac atc aac cag atg ctg      914
112 Thr Ile Lys Ala Ala Ala Cys Thr Pro Tyr Asp Ile Asn Gln Met Leu
113             290                 295                 300
115 ggt acc agc atg acc aag atc ctg aag gag ccc gtg cac ggc gtg aag      962
116 Gly Thr Ser Met Thr Lys Ile Leu Lys Glu Pro Val His Gly Val Lys
117             305                 310                 315
119 gcc gcc cag atc tac cag gag ccc ttc aag aac ctg aag acc ggc gag     1010
120 Ala Ala Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Glu
121             320                 325                 330
123 ccc atc gtg ggc gcc gag acc ttc tac gtg gac ggc gcc gcc aac gtg     1058
124 Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Val
```

RAW SEQUENCE LISTING   DATE: 10/06/2005
PATENT APPLICATION: US/10/550,651   TIME: 11:44:04

Input Set : A:\67675-06 Seq List.txt
Output Set: N:\CRF4\10062005\J550651.raw

```
125      335                340                345
127 atc tac cag tac atg gac gac ctg ctg ctg tgg aag ggc gag ggc gcc      1106
128 Ile Tyr Gln Tyr Met Asp Asp Leu Leu Leu Trp Lys Gly Glu Gly Ala
129 350                355                360                365
131 gtg aag gcc gcc cgc atc cgc acc tgg aag agc ctg gtg aag cac ccc      1154
132 Val Lys Ala Ala Arg Ile Arg Thr Trp Lys Ser Leu Val Lys His Pro
133                370                375                380
135 aag gtg agc agc gag gtg cac atc gcc gtg cgc cac ttc ccc cgc atc      1202
136 Lys Val Ser Ser Glu Val His Ile Ala Val Arg His Phe Pro Arg Ile
137                385                390                395
139 tgg gcc gtg cgc cac ttc ccc cgc ccc tgg gcc atc atc cgc atc ctg      1250
140 Trp Ala Val Arg His Phe Pro Arg Pro Trp Ala Ile Ile Arg Ile Leu
141          400                405                410
143 cag cag ctg aag gcc gcc gtg ggc ttc ccc gtg cgc ccc cag gtg ccc      1298
144 Gln Gln Leu Lys Ala Ala Val Gly Phe Pro Val Arg Pro Gln Val Pro
145          415                420                425
147 ctg cgc ccc atg acc tac aag ggc gcc gtg gac ctg agc cac ttc ctg      1346
148 Leu Arg Pro Met Thr Tyr Lys Gly Ala Val Asp Leu Ser His Phe Leu
149 430                435                440                445
151 aag gag aag ggc ggc ctg ggc ccc ggc gtg cgc tac ccc ctg acc ttc      1394
152 Lys Glu Lys Gly Gly Leu Gly Pro Gly Val Arg Tyr Pro Leu Thr Phe
153                450                455                460
155 ggc tgg tgc tac aag gcc gcc aag acc ctg ccc ctg tgc gtg acc ctg      1442
156 Gly Trp Cys Tyr Lys Ala Ala Lys Thr Leu Pro Leu Cys Val Thr Leu
157          465                470                475
159 acc gtg tac tac ggc gtg ccc gtg tgg aag gag gcc acc acc acc ctg      1490
160 Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu
161          480                485                490
163 cgc gcc atc gag gcc cag cag cac ctg gag cgc tac ctg aag gac ggc      1538
164 Arg Ala Ile Glu Ala Gln Gln His Leu Glu Arg Tyr Leu Lys Asp Gly
165          495                500                505
167 ggc ctg tag ctcgagtagc                                              1557
168 Gly Leu
169 510
```

172 <210> SEQ ID NO: 2
173 <211> LENGTH: 511
174 <212> TYPE: PRT
175 <213> ORGANISM: Artificial Sequence
177 <220> FEATURE:
178 <223> OTHER INFORMATION: Construct encoding polyepitope polypeptide.
180 <400> SEQUENCE: 2

```
182 Met Gln Ile Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu
183 1                5                10                15
186 Val Glu Pro Ser Asp Thr Ile Glu Asn Val Lys Ala Lys Ile Gln Asp
187                20                25                30
190 Lys Glu Gly Ile Pro Pro Asp Gln Gln Arg Leu Ile Phe Ala Gly Lys
191          35                40                45
194 Gln Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn Ile Gln Lys Glu
195          50                55                60
```

```
198 Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Ala Glu Leu Arg Ser
199 65              70              75                  80
202 Leu Tyr Asn Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Lys
203             85                  90                  95
206 Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Trp Ala Ser Arg Glu
207             100                 105             110
210 Leu Glu Arg Phe Lys Ala Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp
211         115                 120             125
214 Val Lys Val Val Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser
215     130                 135             140
218 Ala Leu Ser Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Thr Ser
219 145                 150             155                 160
222 Thr Leu Gln Glu Gln Ile Gly Trp Lys Ala Ala Asn Pro Pro Ile Pro
223             165                 170             175
226 Val Gly Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile
227             180                 185             190
230 Val Arg Met Tyr Ser Pro Thr Ser Ile Phe Arg Asp Tyr Val Asp Arg
231         195                 200             205
234 Phe Tyr Lys Thr Leu Arg Ala Val Gln Asn Ala Asn Pro Asp Cys Lys
235     210                 215             220
238 Thr Ile Leu Lys Ala Leu Ala Cys Gln Gly Val Gly Gly Pro Gly His
239 225                 230             235                 240
242 Lys Lys Ala Ala Ile Thr Leu Trp Gln Arg Pro Leu Val Thr Val Leu
243             245                 250             255
246 Asp Val Gly Asp Ala Tyr Phe Ser Val Trp Lys Gly Ser Pro Ala Ile
247             260                 265             270
250 Phe Gln Ser Lys Leu Arg Gly Pro Gly Arg Ala Phe Val Thr Ile Lys
251         275                 280             285
254 Ala Ala Ala Cys Thr Pro Tyr Asp Ile Asn Gln Met Leu Gly Thr Ser
255     290                 295             300
258 Met Thr Lys Ile Leu Lys Glu Pro Val His Gly Val Lys Ala Ala Gln
259 305                 310             315                 320
262 Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys Thr Gly Glu Pro Ile Val
263             325                 330             335
266 Gly Ala Glu Thr Phe Tyr Val Asp Gly Ala Ala Asn Val Ile Tyr Gln
267             340                 345             350
270 Tyr Met Asp Asp Leu Leu Leu Trp Lys Gly Glu Gly Ala Val Lys Ala
271         355                 360             365
274 Ala Arg Ile Arg Thr Trp Lys Ser Leu Val Lys His Pro Lys Val Ser
275     370                 375             380
278 Ser Glu Val His Ile Ala Val Arg His Phe Pro Arg Ile Trp Ala Val
279 385                 390             395                 400
282 Arg His Phe Pro Arg Pro Trp Ala Ile Ile Arg Ile Leu Gln Gln Leu
283             405                 410             415
286 Lys Ala Ala Val Gly Phe Pro Val Arg Pro Gln Val Pro Leu Arg Pro
287         420                 425             430
290 Met Thr Tyr Lys Gly Ala Val Asp Leu Ser His Phe Leu Lys Glu Lys
291         435                 440             445
294 Gly Gly Leu Gly Pro Gly Val Arg Tyr Pro Leu Thr Phe Gly Trp Cys
```

RAW SEQUENCE LISTING   DATE: 10/06/2005
PATENT APPLICATION: US/10/550,651   TIME: 11:44:04

Input Set : A:\67675-06 Seq List.txt
Output Set: N:\CRF4\10062005\J550651.raw

```
295      450                    455                    460
298 Tyr Lys Ala Ala Lys Thr Leu Pro Leu Cys Val Thr Leu Thr Val Tyr
299 465                    470                    475                    480
302 Tyr Gly Val Pro Val Trp Lys Glu Ala Thr Thr Thr Leu Arg Ala Ile
303                    485                    490                    495
306 Glu Ala Gln Gln His Leu Glu Arg Tyr Leu Lys Asp Gly Gly Leu
307                    500                    505                    510
310 <210> SEQ ID NO: 3
311 <211> LENGTH: 1323
312 <212> TYPE: DNA
313 <213> ORGANISM: Artificial Sequence
315 <220> FEATURE:
316 <223> OTHER INFORMATION: Construct encoding polyepitope polypeptide.
319 <220> FEATURE:
320 <221> NAME/KEY: CDS
321 <222> LOCATION: (7)..(1317)
322 <223> OTHER INFORMATION: Sequence encoding MCMVCTL (no ubiquitin) polyepitope
polypeptide.
324 <400> SEQUENCE: 3
325 gctagc atg gag ctc cgc tcc ctc tac aac acc gtg gcc acc ctc tac          48
326        Met Glu Leu Arg Ser Leu Tyr Asn Thr Val Ala Thr Leu Tyr
327         1               5                    10
329 tgc gtg cac cag cgc atc aag atc cgc ctg cgc ccc ggc ggc aag aag          96
330 Cys Val His Gln Arg Ile Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys
331 15                    20                    25                    30
333 aag tac tgg gcc agc cgc gag ctg gag cgc ttc aag gcc gcc atc agc         144
334 Lys Tyr Trp Ala Ser Arg Glu Leu Glu Arg Phe Lys Ala Ala Ile Ser
335                    35                    40                    45
337 ccc cgc acc ctg aac gcc tgg gtg aag gtg gtg aag gcc ttc agc ccc         192
338 Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Lys Ala Phe Ser Pro
339                    50                    55                    60
341 gag gtg atc ccc atg ttc agc gcc ctg agc gag ggc gcc acc ccc cag         240
342 Glu Val Ile Pro Met Phe Ser Ala Leu Ser Glu Gly Ala Thr Pro Gln
343             65                    70                    75
345 gac ctg aac acc atg acc agc acc ctg cag gag cag atc ggc tgg aag         288
346 Asp Leu Asn Thr Met Thr Ser Thr Leu Gln Glu Gln Ile Gly Trp Lys
347         80                    85                    90
349 gcc gcc aac ccc ccc atc ccc gtg ggc gac atc tac aag cgc tgg atc         336
350 Ala Ala Asn Pro Pro Ile Pro Val Gly Asp Ile Tyr Lys Arg Trp Ile
351 95                    100                    105                    110
353 atc ctg ggc ctg aac aag atc gtg cgc atg tac agc ccc acc agc atc         384
354 Ile Leu Gly Leu Asn Lys Ile Val Arg Met Tyr Ser Pro Thr Ser Ile
355                    115                    120                    125
357 ttc cgc gac tac gtg gac cgc ttc tac aag acc ctg cgc gcc gtg cag         432
358 Phe Arg Asp Tyr Val Asp Arg Phe Tyr Lys Thr Leu Arg Ala Val Gln
359                    130                    135                    140
361 aac gcc aac ccc gac tgc aag acc atc ctg aag gcc ctg gcc tgc cag         480
362 Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Lys Ala Leu Ala Cys Gln
363             145                    150                    155
365 ggc gtg ggc ggc ccc ggc cac aag aag gcc gcc atc acc ctg tgg cag         528
```

VERIFICATION SUMMARY   DATE: 10/06/2005
PATENT APPLICATION: **US/10/550,651**   TIME: 11:44:05

Input Set : **A:\67675-06 Seq List.txt**
Output Set: **N:\CRF4\10062005\J550651.raw**

L:14 M:270 C: Current Application Number differs, Replaced Current Application No
L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date